

Comments on the article titled “The regulation of AI liability in Europe: a critical overview of two recent Directive Proposals - the (new) ALD and the (revised) PLD” by Beatriz Garcia

Mariana de Lemos Campos

Vol. 11 No. 3
Dec 2024
e-publica.pt

ISSN 2183-184x

Com o apoio de:

fct Fundação
para a Ciência
e a Tecnologia

COMMENTS ON THE ARTICLE TITLED “THE REGULATION OF AI LIABILITY IN EUROPE: A CRITICAL OVERVIEW OF TWO RECENT DIRECTIVE PROPOSALS - THE (NEW) AILD AND THE (REVISED) PLD” BY BEATRIZ GARCIA

COMENTÁRIOS SOBRE O ARTIGO INTITULADO “A REGULAÇÃO DA RESPONSABILIDADE PELA UTILIZAÇÃO DA IA: COMENTÁRIO CRÍTICO A DUAS RECENTES PROPOSTAS DE DIRETIVA: A (NOVA) DRIA E A (REVISTA) DRP”, DE BEATRIZ GARCIA

MARIANA DE LEMOS CAMPOS¹

Faculdade de Direito da Universidade de Lisboa

Portugal

marianalemosc@gmail.com

Abstract: This commentary expands on Beatriz Garcia’s critique of the Artificial Intelligence Liability Directive (AILD) and the revised Product Liability Directive (PLD), focusing on their interaction with the recently enacted AI Act. Garcia identifies key issues: the PLD’s narrow focus on software as a “product” and the AILD’s burdensome evidentiary requirements for claimants seeking AI-related compensation. The AI Act, effective August 2024, establishes a risk-based regulatory framework prioritizing harm prevention over liability. While fostering innovation through regulatory sandboxes, it lacks civil compensation mechanisms, creating regulatory gaps. Case studies—an autonomous vehicle fatality and discriminatory AI hiring practices—highlight the difficulties of applying AILD and PLD frameworks, particularly in addressing non-physical harm and the “black box” nature of AI systems. This analysis underscores a disconnect between the AI Act’s harm-prevention focus and the liability frameworks, favoring companies while complicating compensation for individuals. A recalibrated regulatory approach is needed to align innovation with accountability.

Keywords: AI liability, AI Act, AILD, PLD, regulatory framework

Resumo: Este comentário expande a crítica de Beatriz Garcia às Diretivas sobre Responsabilidade da Inteligência Artificial (DRIA) e pelos Produtos (DRP), explorando sua interação com o novo AI Act. Garcia aponta problemas: o foco restrito da DRP em software como “produto” e os elevados encargos probatórios da DRIA. O AI Act, em vigor desde agosto de 2024, prioriza a prevenção de danos através de um quadro regulatório baseado no risco, promovendo inovação, mas sem mecanismos de compensação civil, criando lacunas regulatórias. Estudos de caso—uma fatalidade com um veículo autónomo e práticas discriminatórias de recrutamento por IA—revelam dificuldades na aplicação das diretivas, especialmente em danos não físicos e a opacidade dos sistemas de IA. A análise destaca o desalinhamento entre o foco preventivo do AI Act e os quadros de responsabilidade, favorecendo empresas e dificultando a compensação de indivíduos. Uma abordagem regulatória ajustada é necessária para equilibrar inovação e responsabilidade.

Palavras-chave: responsabilidade da IA, AI Act, DRIA, DRP, quadro regulatório

1. PhD candidate at the University of Lisbon School of Law.

The author presents a compelling analysis, focusing on criticisms of two Directive Proposals. The conclusion can be summarized as follows: the primary issue with the Product Liability Directive (PLD) lies in defining software as a “product,” since most problems arise from the normal functioning of software rather than inherent product defects. Meanwhile, the Artificial Intelligence Liability Directive (AILD) imposes significant challenges on natural persons seeking to prove fault and causality, where a rebuttable presumption for both elements should have been established.

The author also notes that a review of both Directive Proposals will likely be necessary following the implementation of the AI Act and ends the article mentioning that “with the recent approval of the AI, we are curious to witness what will happen to both proposals, specially the AILD”.

Following this sentence, we seek to begin our analysis from this latest point: the AI Act.

Our aim is to dive into the arguments presented in the article and establish the connection between this new regulatory framework with the PLD and the AILD, to explore how they may interact in practice.

For this purpose, this comment will be divided into three parts: first, a general overview of the AI Act; second, two case studies will be presented to examine the possible applicable liability and the associated challenges; and finally, a conclusion with key remarks will be provided.

Turning first to the AI Act, it was officially published on July 12 of this year (2024), coming into force in August, although additional *vacatio legis* periods and deadlines for specific obligations remain.

Its primary objective is to regulate the use of artificial intelligence within the European Union. Unlike the AILD and the reformed PLD, the AI Act does not focus directly on liability or compensation mechanisms; rather, its emphasis is on preventing harm.

The AI Act's core objective is to establish a risk-based regulatory framework for AI systems, empowering regulatory bodies to impose fines for non-compliance. However, as mentioned, it does not address civil compensation.

To operationalize this risk-based approach, the AI Act categorizes AI systems based on four levels of risk: minimal risk, transparency risk, high risk, and unacceptable risk.

Unacceptable risks, which are defined in Article 5, are prohibited. These include AI applications that could infringe upon Union law or values, such as “social scoring”.

High-risk AI systems, as defined in Article 6, are the ones that may affect individual health, safety, or fundamental rights and therefore must meet strict regulatory requirements in order to operate. These include risk mitigation measures, user transparency, human oversight, record-keeping, and technical documentation. An example of high-risk AI would be AI system used in recruitment processes.

Transparency risks arise where there is a potential for manipulation, requiring users to be informed when they are interacting with AI, as in the case of chatbots.

Minimal risk AI systems, such as AI-enabled video games, are not subject to specific obligations but may voluntarily comply with guidelines under the AI Act.

In this sense, it's clear that a significant portion of artificial intelligence systems, particularly those exhibiting increased complexity, are likely to be classified as high-risk under the applicable regulatory framework.

On another note, a notable aspect of the AI Act is its explicit aim to support innovation. This is reflected in Chapter VI (entirely dedicated to this aim), which promotes regulatory sandboxes (Article 57), real-world testing under specific conditions, and measures for AI providers and deployers (Article 62).

This innovation-friendly approach seeks to address the chilling effect that liability concerns have had on European companies' adoption of AI technologies.

As seen, to address this type of civil liability in Europe Union, the AILD and PLD are designed; AILD theoretically ensuring that victims have the same level of protection in cases involving AI systems as they would in non-AI contexts, following the principle of non-discrimination.

Two case studies will now be presented to elucidate the practical implications of the aforementioned Directive Proposals.

The first case is that of Elaine Herzberg, who was killed by an autonomous Uber vehicle in Arizona in 2018.

Both the National Transportation Safety Board (NTSB) and the Tempe Police Department conducted exhaustive investigations, requiring approximately 20 months to arrive at their conclusions. These findings highlighted critical shortcomings in the AI system, including its failure to accurately classify the pedestrian as such - misidentifying her instead as a bicycle and other objects - and the operator's inability to assume control of the autonomous vehicle and engage the braking system in time to prevent the fatal incident.

In a hypothetical scenario where a similar accident was to occur in the present day within the jurisdiction of European Union, a claimant would face formidable challenges in establishing plausibility, if the AILD were applicable. The inherent complexity and opacity of the AI system, often termed the "black box effect", would impose a considerable burden on the claimant, requiring them to demonstrate non-compliance, fault, and causation. These demands would reflect the many difficulties encountered in this case, underscoring the practical obstacles of securing liability under the AILD framework.

A similar challenge would arise under the PLD, given that several months were required to ascertain whether the autonomous vehicle itself was defective or if the incident was solely attributable to the operator's inattentive behavior.

If establishing these elements proved onerous for entities with significant resources, technical expertise, and investigatory authority, such as the NTSB and Tempe Police Department, it is reasonable to surmise that an individual claimant, lacking such institutional support, would confront an even more prohibitive degree of difficulty in advancing a claim under the current regulatory framework.

The second example concerns discriminatory hiring practices by an AI system.

To present this example, we will use a fictional scenario presented by Phil Lee (2024). Assuming a society where all people are either green or blue,

and woman or man, an entity has the aim to achieve equal representation by hiring 50% of each color (blue and green) and 50% of each gender (men and women). However, the AI algorithm could associate "blue" with men and "green" with women, and it would disproportionately favor "blue men" and "green women", leading to discrimination against "blue women" and "green men". In this case, the algorithm ensured the top-level fairness, but disregarded the subgroups, leading to discrimination.

This simple example illustrates how biased data can lead to unfair outcomes, despite well-meaning intentions.

Continuing with the hypothetical scenario and considering a candidate who has been subjected to this discriminatory practice, with evident harm, this candidate would be entitled to seek compensation through the civil liability mechanisms available against the company.

However, notwithstanding the evident damage suffered, the candidate would be unable to invoke strict liability under the PLD because the harm in question (discrimination) does not constitute a physical injury, which is the only type of damage that the PLD addresses.

Both examples illustrate the considerable difficulties inherent in applying AILD and PLD in a manner that effectively holds companies liable for harm resulting from artificial intelligence systems. These challenges raise fundamental questions regarding the adequacy of the proposed regulatory framework, particularly in ensuring that victims can seek redress and obtain appropriate compensation for the damages.

While the AI Act focuses on preventing harm by imposing strict compliance requirements on entities, the subsequent frameworks established by the AILD and PLD seem to facilitate the operational capacity of companies at the expense of the individual's ability to obtain civil compensation, given the evidentiary challenges that individuals must overcome.

This misalignment results in an environment where, despite meeting compliance obligations under the AI Act, companies may still evade liability for damages caused by their AI systems.

In conclusion, while the European Union is appropriately focused on fostering innovation within the AI sector, this approach inadvertently creates significant obstacles for natural persons seeking to assert their rights in relation to AI-related damage.

The strict requirements of the AI Act, designed to prevent harm, stand in stark contrast to the provisions of the AILD and PLD, which appear to ease the burdens on companies while complicating it for individuals. This imbalance necessitates a thorough reassessment of the existing regulatory landscape to ensure the establishment of an integrated and effective legal framework governing artificial intelligence technologies.

References

Lee P. Why is algorithmic design so important to protect against discrimination? LinkedIn post. September 2024. Available at: https://www.linkedin.com/posts/phillee77_why-is-algorithmic-design-so-important-to-activity-7236667283312381953-1Ely?utm_source=share&utm_medium=member_desktop